

Interpolation and Quadrature

Abigail Poleway

Faculty Mentor: Dr. Elliott Bertrand

MA 398: Senior Seminar

Instructor: Dr. Elizabeth Tripp

Fall 2023

Abstract

In this paper, we will first examine polynomial interpolation. Interpolation allows us to construct a polynomial through a given set of data points. Interpolation is also used to develop the Newton-Cotes formulas, which are methods for numerical integration, or quadrature, that can approximate more complicated definite integrals. However, there is a way to arrive at even more precise approximations, and we will discover this through a discussion of Gaussian Quadrature.

1 Introduction

In this paper, we will explore the useful applications of polynomial interpolation and quadrature, or numerical integration, methods. Interpolation is a process we can use to construct polynomials through a set of distinct points, or nodes. This process allows us to choose points from any given function, including those that are difficult and involved, to construct a polynomial approximation that is much simpler to use. Although interpolation is an ancient practice, its widespread use expanded through the shipbuilding and aircraft industries, aerodynamic studies of automobiles, and computer typesetting [3]. Interpolation helps us organize and consolidate data in more efficient ways so that we may enhance our comprehension of various mathematical and scientific enigmas.

Elementary functions are “polynomials, rational functions, power functions, exponential functions, logarithmic functions, trigonometric and inverse trigonometric functions, hyperbolic and inverse hyperbolic functions, and all functions that can be” [5] constructed using finitely many operations; these five operations include “addition, subtraction, multiplication, division, and composition” [5]. There are numerous integrals that are considered impossible to solve exactly in closed form based on their complexity, as they do not have antiderivatives that are elementary functions. So, we must utilize other methods to help us compute them; we call these numerical integration methods. As there is a certain degree of error present in numerical integration approximations, it is important that we acknowledge that there may be more accurate, effective methods than others, and this is where we will look into Gaussian Quadrature. This method is beneficial to the analysis of data sets and is particularly

famous for its applications found throughout probability and statistics, which is why it is an advantageous topic to investigate.

2 Interpolation

In this paper, we will observe how the process of interpolation helps us determine functions that pass through certain data points. This method can help us model a variety of situations, which provides opportunities to compute approximations and make useful predictions. We will begin by defining interpolation.

Definition. A function $y = F(x)$ ***interpolates*** the points $(x_1, y_1), \dots, (x_n, y_n)$ if $F(x_k) = y_k$ for each $1 \leq k \leq n$.

In other words, given a set of n nodes, or points, we can use interpolation to construct a polynomial that passes through them. Recall that F is only a function if each x coordinate is mapped to exactly one y coordinate; the x_n 's are unique, so a function may successfully pass through them. As long as each x coordinate is different, no matter how many there are, there will always be a polynomial that goes through each point.

It is useful to think of interpolation as the opposite of evaluation [3]. In order to evaluate a given polynomial, we compute values using the available information to determine points that are located on the curve. On the other hand, in order to interpolate a polynomial, we are given a set of data points, and we must determine a polynomial that produces them.

2.1 Lagrange Interpolation

Given a set of n data points, we can find the interpolating polynomial using an explicit formula. This formula, called the Lagrange interpolating formula, allows us to construct polynomials of degree at most $d = n - 1$ that travel through the set of n given points [3]. The polynomial that is constructed for each set of points is the ***Lagrange interpolating polynomial*** for these designated points, which we define next.

Definition. Consider n distinct **nodes** (points) $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$.

The ***Lagrange interpolating polynomial*** of degree at most $n - 1$ is

$$P_{n-1}(x) = \sum_{k=1}^n y_k L_k(x),$$

where for each $k = 1, \dots, n$, we define the function L_k as follows:

$$L_k(x) = \frac{(x - x_1) \cdots (x - x_{k-1})(x - x_{k+1}) \cdots (x - x_n)}{(x_k - x_1) \cdots (x_k - x_{k-1})(x_k - x_{k+1}) \cdots (x_k - x_n)} = \prod_{\substack{i=1 \\ i \neq k}}^n \frac{(x - x_i)}{(x_k - x_i)}. \quad (1)$$

By design, $L_k(x_k) = 1$ while $L_k(x_j) = 0$ for all $j \neq k$.

We will implement this formula to observe its effects using an example.

Example 2.1. Suppose we have the points $(x_1, y_1), (x_2, y_2), (x_3, y_3)$. Then the Lagrange interpolating polynomial for these points is

$$P_2(x) = y_1 \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)} + y_2 \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)} + y_3 \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}.$$

Note that this is a degree 2 or less interpolating polynomial because $d = n - 1$ and $n = 3$, so it follows that $d = 3 - 1 = 2$.

We will look at an example of this degree 2 interpolating polynomial by plugging in actual points.

Example 2.2. *Find the interpolating polynomial through the points $(0, 2)$, $(1, 3)$, and $(4, 2)$.*

Implementing the Lagrange interpolating formula,

$$\begin{aligned} P_2(x) &= 2 \frac{(x-1)(x-4)}{(0-1)(0-4)} + 3 \frac{(x-0)(x-4)}{(1-0)(1-4)} + 2 \frac{(x-0)(x-1)}{(4-0)(4-1)} \\ &= \frac{1}{2}(x^2 - 5x + 4) - (x^2 - 4x) + \frac{1}{6}(x^2 - x) \\ &= -\frac{1}{3}x^2 + \frac{4}{3}x + 2. \end{aligned}$$

We can check that this polynomial truly interpolates these points by plugging each x and y into their respective places in $P_2(x)$. So,

$$P_2(0) = -\frac{1}{3}(0)^2 + \frac{4}{3}(0) + 2 = 2,$$

$$P_2(1) = -\frac{1}{3}(1)^2 + \frac{4}{3}(1) + 2 = 3,$$

$$P_2(4) = -\frac{1}{3}(4)^2 + \frac{4}{3}(4) + 2 = 2.$$

Thus, $P_2(x)$ is indeed the interpolating polynomial through the points $(0, 2)$, $(1, 3)$, $(4, 2)$.

Next, we must consider an important theorem concerning the concept of polynomial interpolation.

Theorem 2.3. Main Theorem of Polynomial Interpolation. *Let $(x_1, y_1), \dots, (x_n, y_n)$ be n points in the plane with distinct x_i . Then, there exists one and only one polynomial P of degree $n - 1$ or less that satisfies $P(x_i) = y_i$ for $i = 1, \dots, n$.*

This theorem provides uniqueness for polynomials of degree $n - 1$ or less. To demonstrate why this must be true, consider the following examples. Given 4 points, this polynomial could have a degree of 3, 2, 1, or 0 since $4 - 1 = 3$ and the degree of the polynomial can only be 3 or less. These 4 nodes could be colinear, making this a polynomial of degree 1 or 0, or a straight line; they could form a parabola, making it a polynomial of degree 2; or it could be a cubic polynomial, making it a polynomial of degree 3. This is why we say degree $n - 1$ or less.

2.2 Interpolation Error

Since interpolation is not perfect, we must analyze the error produced when using this strategy. So, we will define interpolation error.

Definition. *Given the original function $y = f(x)$ and the interpolating polynomial $P(x)$ that was constructed using its points, the **interpolation error** at x is $f(x) - P(x)$.*

In other words, if we subtract the interpolating polynomial, evaluated at x , from the function that supplied us with the data points we used to create the interpolating polynomial, we can gauge how much error is present between them.

Since interpolation error cannot be calculated exactly, we can use the following theorem to obtain error bounds, which will tell us approximately how large the present error is [3].

Theorem 2.4. *Assume that $P(x)$ is the degree $n - 1$ or less interpolating polynomial fitting the n points $(x_1, y_1), \dots, (x_n, y_n)$. The interpolation error is*

$$f(x) - P(x) = \frac{(x - x_1)(x - x_2) \dots (x - x_n)}{n!} f^{(n)}(c),$$

where c lies between the smallest and largest of the numbers x, x_1, \dots, x_n .

For more information on interpolation error, see [3].

3 Numerical Integration and Quadrature

Given a function f on an interval $[a, b]$, we can construct an interpolating polynomial through $f(x)$ at certain points. Since evaluating a definite integral of a polynomial is relatively simple, this strategy can be used to approximate $\int_a^b f(x)$. Integration is important in many mathematical fields, like probability and statistics, and individuals working in these fields use particular integrals to draw conclusions about the data they analyze. The methods they decide to use to compute values have an impact on these conclusions depending on how accurate they are. Some examples, like the Newton-Cotes formulas, are quadrature methods built on interpolating polynomials that use equally spaced points to make approximations.

3.1 Newton-Cotes Formulas

The Newton-Cotes methods of approximating integrals, based upon interpolation, include the formulas for the Trapezoid Rule and Simpson's Rule. We will be able to see how accurate they are based on their degree of precision. First, we analyze the Trapezoid Rule.

3.2 Trapezoid Rule

We can use the “simplest application of interpolation-based integration” [3], the Trapezoid Rule, when $n = 1$. In other words, the Trapezoid Rule uses two nodes to help us

construct a degree 1 interpolating polynomial, which are the endpoints of the interval in question [3]. In order to make approximations, this rule approximates definite integrals by finding the area underneath the line that connects these endpoints.

For instance, let $P(x)$ be a degree 1 interpolating polynomial through the points (x_1, y_1) and (x_2, y_2) . If we implement the Lagrange interpolating polynomial formula to find the polynomial for this case, then add its error term to maintain as much accuracy as possible, we obtain

$$f(x) = y_1 \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_1}{x_2 - x_1} + \frac{(x - x_1)(x - x_2)}{2!} f''(c) = P(x) + E(x).$$

When we integrate both sides of this equation, we have

$$\int_{x_1}^{x_2} f(x) dx = \int_{x_1}^{x_2} P(x) dx + \int_{x_1}^{x_2} E(x) dx.$$

It follows that

$$\begin{aligned} \int_{x_1}^{x_2} P(x) dx &= y_1 \int_{x_1}^{x_2} \frac{x - x_2}{x_1 - x_2} dx + y_2 \int_{x_1}^{x_2} \frac{x - x_1}{x_2 - x_1} dx \\ &= \frac{y_1}{x_1 - x_2} \left(\frac{x^2}{2} - x_2 x \right) \Big|_{x_1}^{x_2} + \frac{y_2}{x_2 - x_1} \left(\frac{x^2}{2} - x_1 x \right) \Big|_{x_1}^{x_2} \\ &= \frac{y_1}{x_1 - x_2} \left(\left(\frac{x_2^2}{2} - x_2^2 \right) - \left(\frac{x_1^2}{2} - x_2 x_1 \right) \right) + \frac{y_2}{x_2 - x_1} \left(\left(\frac{x_2^2}{2} - x_1 x_2 \right) - \left(\frac{x_1^2}{2} - x_1^2 \right) \right) \\ &= \frac{1}{x_2 - x_1} \left(-y_1 \left(-\frac{x_2^2}{2} - \frac{x_1^2}{2} + x_1 x_2 \right) + y_2 \left(\frac{x_2^2}{2} - x_1 x_2 + \frac{x_1^2}{2} \right) \right) \\ &= \frac{1}{x_2 - x_1} \left(y_1 \left(\frac{x_2^2}{2} - x_1 x_2 + \frac{x_1^2}{2} \right) + y_2 \left(\frac{x_2^2}{2} - x_1 x_2 + \frac{x_1^2}{2} \right) \right) \\ &= \frac{1}{x_2 - x_1} \left(\frac{y_1}{2} (x_2^2 - 2x_1 x_2 + x_1^2) + \frac{y_2}{2} (x_2^2 - 2x_1 x_2 + x_1^2) \right) \\ &= \frac{1}{x_2 - x_1} \left(\left(\frac{y_1}{2} + \frac{y_2}{2} \right) (x_2 - x_1)^2 \right). \end{aligned}$$

Further simplification yields

$$\begin{aligned} &= (x_2 - x_1) \left(\frac{y_1 + y_2}{2} \right) \\ &= y_1 \frac{h}{2} + y_2 \frac{h}{2} \\ &= h \frac{y_1 + y_2}{2}, \end{aligned}$$

where $h = x_2 - x_1$ is the length of the interval.

So,

$$\int_{x_1}^{x_2} P(x) dx = \frac{h}{2}(y_1 + y_2).$$

Note that

$$\int_{x_1}^{x_2} E(x) dx = -\frac{h^3}{12} f''(c).$$

See [3] for the calculations of the error integral.

Thus,

Theorem 3.1. *Trapezoid Rule.*

$$\int_{x_1}^{x_2} f(x) dx = \frac{h}{2}(y_1 + y_2) - \frac{h^3}{12} f''(c),$$

where $h = x_2 - x_1$ and c is between x_1 and x_2 .

Note that h represents the distance between the nodes. Since the Trapezoid Rule's error formula contains a second derivative, the Trapezoid Rule tells us that if $f(x)$ is itself a polynomial of degree 1 or less, the interpolation error is 0, meaning this polynomial is integrated exactly [3]. This is where we observe the degree of precision, which we define below.

Definition. The *degree of precision* of a numerical integration method is the greatest integer k for which all degree k or less polynomials are integrated exactly by the method.

We can observe this result using an example of a polynomial with degree 1.

Example 3.2. Apply the Trapezoid Rule to approximate

$$\int_0^1 (x - 2)dx.$$

Here $(x_1, y_1) = (0, -2)$ and $(x_2, y_2) = (1, -1)$. Using the Trapezoid Rule, we have

$$\begin{aligned}\int_0^1 (x - 2)dx &= \frac{h}{2}(y_1 + y_2) \\ &= \frac{1}{2}(-1 - 2) \\ &= -\frac{3}{2}.\end{aligned}$$

Now, we can integrate this function using traditional calculus methods:

$$\begin{aligned}\int_0^1 (x - 2)dx &= \left(\frac{x^2}{2} - 2x \right) \Big|_0^1 \\ &= \left(\frac{(1)^2}{2} - 2(1) \right) - 0 \\ &= \frac{1}{2} - 2 \\ &= -\frac{3}{2}.\end{aligned}$$

Hence, this example illustrates the Trapezoid Rule having a degree of precision of 1.

3.3 Simpson's Rule

Simpson's Rule can be used when $n = 2$, or when there are three nodes, which are two endpoints and a midpoint [3]. Now, let $P(x)$ be a degree 2 interpolating polynomial through

the points (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) . Again, using the Lagrange interpolating polynomial and adding the error term, we acquire

$$f(x) = y_1 \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} + y_2 \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} + y_3 \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)} + \frac{(x-x_1)(x-x_2)(x-x_3)}{3!} f'''(c).$$

In a similar fashion to the Trapezoid Rule, by integrating both sides of this equation we have what is known as Simpson's Rule.

Theorem 3.3. *Simpson's Rule*

$$\int_{x_1}^{x_3} f(x) dx = \frac{h}{3}(y_1 + 4y_2 + y_3) - \frac{h^5}{90} f^{(iv)}(c),$$

where $h = x_3 - x_2 = x_2 - x_1$ and c is between x_1 and x_3 .

Recall that h represents the distance between nodes. In this case, we can clearly see that they are evenly spaced since $x_3 - x_2 = x_2 - x_1$. The degree of precision for Simpson's Rule is 3, which means that, by definition, it can be used to evaluate the integrals of cubic, quadratic, and linear functions exactly.

We will observe this result using an example of a cubic polynomial.

Example 3.4. *Apply Simpson's Rule to approximate*

$$\int_0^2 (x^3 + 2x^2 - 1) dx.$$

Here $(x_1, y_1) = (0, -1)$, $(x_2, y_2) = (1, 2)$, and $(x_3, y_3) = (2, 15)$. Using Simpson's Rule we have

$$\begin{aligned} \int_0^2 (x^3 + 2x^2 - 1) dx &= \frac{h}{3}(y_1 + 4y_2 + y_3) \\ &= \frac{1}{3}(-1 + 8 + 15) \\ &= \frac{22}{3}. \end{aligned}$$

Now, we will integrate this function using traditional calculus methods.

$$\begin{aligned}\int_0^2 (x^3 + 2x^2 - 1)dx &= \left. \frac{x^4}{4} + \frac{2x^3}{3} - x \right|_0^2 \\ &= \left[\frac{(2)^4}{4} + \frac{2(2)^3}{3} - 2 \right] - 0 \\ &= \frac{16}{4} + \frac{16}{3} - 2 \\ &= \frac{22}{3}.\end{aligned}$$

This example illustrates Simpson's Rule, having a degree of precision of 3. It follows that that the more nodes we look at, the more accurate our approximations are, which is why the Newton-Cotes method of integration is so valuable.

To conclude, we will state a theorem that provides the degree of precision for Newton-Cotes methods.

Theorem 3.5. *Newton-Cotes methods of degree n have a degree of precision n when n is odd. Methods of degree n have a degree of precision $n + 1$ when n is even.*

We can observe this theorem through the Trapezoid Rule and Simpson's Rule. For the Trapezoid Rule, $n=1$, which is odd, and it has a degree of precision of 1. For Simpson's Rule, $n=2$, which is even, and it has a degree of precision of 3.

3.4 Gaussian Quadrature

Recall that Newton-Cotes methods consider evenly spaced points. We must consider whether these methods suffice or if there are other methods that can be used to increase the

accuracy of our approximations. We will analyze the degree of precision in other methods, namely Gaussian Quadrature.

Gaussian Quadrature allows us to approximate definite integrals that otherwise cannot be solved. For example,

$$\int_{-1}^1 e^{-x^2/2} dx$$

is an integral that does not have an elementary antiderivative. It would be helpful to solve this integral because it is implemented in a variety of situations, especially probability and statistics. The Gaussian distribution, also known as the normal distribution or the bell curve, is used in probability and statistics to represent random variables that have unknown distributions and provides data that helps us draw conclusions about the mean and standard deviation of a data set. This particular function, $e^{-x^2/2}$, is a component of the probability density function, which can be used to provide a likelihood that the random variable's value is the same as the sample [7]. Using Gaussian Quadrature, we can approximate this function using an integral so that computations in other disciplines become more accurate and efficient. However, before we define Gaussian Quadrature, we must first investigate the Legendre polynomials.

Before defining the Legendre polynomials, we must consider the property of orthogonality as it is a significant factor in reducing approximation error prompted by interpolation.

Definition. A set of nonzero functions $\{p_0, \dots, p_n\}$ on the interval $[a, b]$ is **orthogonal** on $[a, b]$ if

$$\int_a^b p_j(x)p_k(x)dx = 0$$

whenever $j \neq k$.

To summarize, a set of functions is orthogonal if, when the product of any two distinct functions from the set is integrated over $[a, b]$, the outcome is 0.

We must also consider an important theorem that allows us to implement Gaussian Quadrature.

Theorem 3.6. *If p_0, p_1, \dots, p_n is an orthogonal set of polynomials on the interval $[a, b]$, where $\deg p_i = i$, then p_0, p_1, \dots, p_n is a basis for the vector space of degree at most n polynomials on $[a, b]$.*

We may now define the Legendre polynomials.

Definition. Legendre Polynomials *The Legendre polynomial of degree i is given by*

$$p_i(x) = \frac{1}{2^i i!} \frac{d^i}{dx^i} [(x^2 - 1)^i].$$

We can implement this definition to produce an example.

Example 3.7.

$$\begin{aligned} p_3(x) &= \frac{1}{2^3 3!} \frac{d^3}{dx^3} [(x^2 - 1)^3] \\ &= \frac{1}{(8)(6)} \frac{d^3}{dx^3} [(x^6 - 3x^4 + 3x^2 - 1)] \\ &= \frac{1}{48} (120x^3 - 72x) \\ &= \frac{5}{2}x^3 - \frac{3}{2} \\ p_3(x) &= \frac{1}{2}(5x^3 - 3x). \end{aligned}$$

i	$P_i(x)$
0	1
1	x
2	$1/2(3x^2 - 1)$
3	$1/2(5x^3 - 3x)$
4	$1/8(35x^4 - 30x^2 + 3)$
5	$1/8(63x^5 - 70x^3 + 15x)$
6	$1/16(231x^6 - 315x^4 + 105x^2 - 5)$
7	$1/16(429x^7 - 693x^5 + 315x^3 - 35x)$

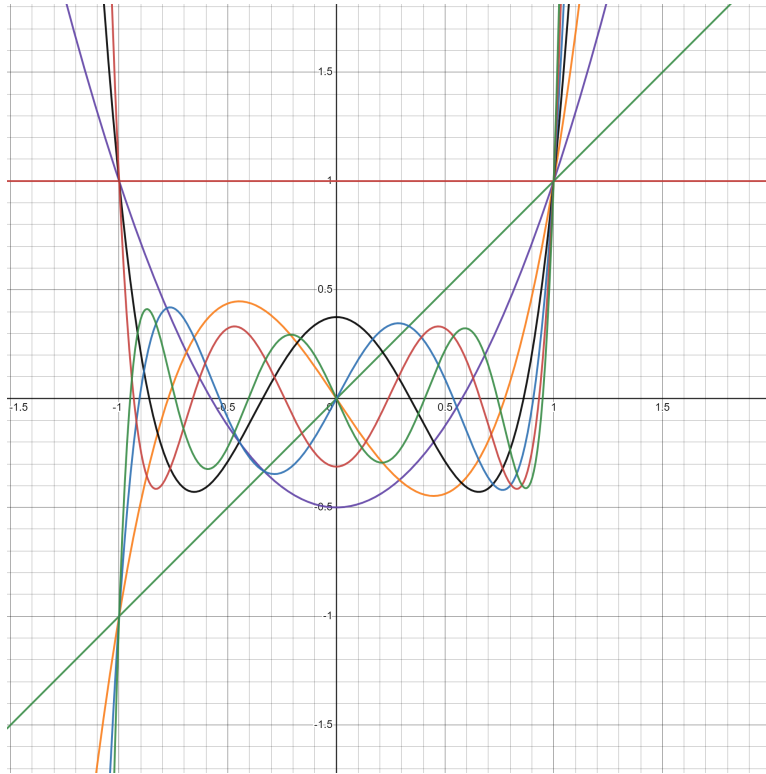
Table 1: The first eight Legendre polynomials.

Thus, the degree 3 Legendre polynomial is

$$p_3(x) = \frac{1}{2}(5x^3 - 3x).$$

The Legendre polynomials are a system of orthogonal polynomials that produce the distinct roots, x_i , which we will use [3]. Gaussian Quadrature can be described as a linear combination of function values of a given function, the integrand of our interpolated polynomial, evaluated at Legendre roots and multiplied by coefficients, c_i [3].

Included below are the first eight Legendre polynomials based on our given definition.



Now, we may define Gaussian Quadrature.

Definition. *Gaussian Quadrature*

$$\int_{-1}^1 f(x)dx \approx \sum_{i=1}^n c_i f(x_i)$$

where

$$c_i = \int_{-1}^1 L_k(x)dx, k = 1, \dots, n.$$

Recall the functions $L_k(x)$ are defined in Equation (1).

In Gaussian Quadrature, the spacing between the evaluation points is chosen optimally instead of equally, with the goal of minimizing the approximation error, which is why Gaus-

n	roots x_i	coefficients c_i
2	$-\sqrt{1/3} = -0.57735026918963$ $\sqrt{1/3} = 0.57735026918963$	$1 = 1.000000000000000$ $1 = 1.000000000000000$
3	$-\sqrt{3/5} = -0.77459666924148$ $0 = 0.000000000000000$ $\sqrt{3/5} = 0.77459666924148$	$\frac{5}{9} = 0.555555555555555$ $\frac{8}{9} = 0.888888888888888$ $\frac{5}{9} = 0.555555555555555$
4	$-\sqrt{\frac{15 + 2\sqrt{30}}{35}} = -0.86113631159405$ $-\sqrt{\frac{15 - 2\sqrt{30}}{35}} = -0.33998104358486$ $\sqrt{\frac{15 - 2\sqrt{30}}{35}} = 0.33998104358486$ $\sqrt{\frac{15 + 2\sqrt{30}}{35}} = 0.86113631159405$	$\frac{90 - 5\sqrt{30}}{180} = 0.34785484513745$ $\frac{90 + 5\sqrt{30}}{180} = 0.65214515486255$ $\frac{90 + 5\sqrt{30}}{180} = 0.65214515486255$ $\frac{90 - 5\sqrt{30}}{180} = 0.34785484513745$

Table 2: Gaussian Quadrature coefficients, where x_i represents the roots of the n th Legendre polynomials, and c_i represents the coefficients for Gaussian Quadrature.

Gaussian Quadrature approximations are much more accurate and useful compared to other integration techniques. The interval $[-1, 1]$ produces the nodes that we can evaluate and the chosen coefficients are selected such that this approximation error remains relatively small. We will examine the definition of Gaussian Quadrature through an example.

Example 3.8. *Using Gaussian Quadrature, approximate*

$$\int_{-1}^1 e^{-x^2/2} dx.$$

Let $f(x) = e^{-x^2/2}$. First, using $n = 2$, we have

$$\int_{-1}^1 e^{-x^2/2} dx \approx c_1 f(x_1) + c_2 f(x_2)$$

$$= 1f(-\sqrt{1/3}) + 1f(\sqrt{1/3})$$

$$\approx 1.69296344978123.$$

Now, using $n = 3$, we have an approximation of

$$\frac{5}{9}f(-\sqrt{3/5}) + \frac{8}{9}f(0) + \frac{5}{9}f(\sqrt{3/5})$$

$$\approx 1.71202024520191.$$

Finally, when $n = 4$ it follows that

$$c_1f(x_1) + c_2f(x_2) + c_3f(x_3) + c_4f(x_4)$$

$$\approx 1.71122450459949.$$

Note that the correct answer is 1.71124878378430 [3]. We can see that as n increases, the accuracy of our approximation increases as well. Now, we must consider the degree of precision for Gaussian Quadrature, which is given in the theorem below.

Theorem 3.9. *The Gaussian Quadrature method, using the degree n Legendre polynomial on $[-1, 1]$, has a degree of precision $2n - 1$.*

For following proof, see [1], [3], and [6] for reference.

Proof. Let $S(x)$ be any polynomial of degree $2n - 1$. We will show that it can be integrated exactly using Gaussian Quadrature. By polynomial long division,

$$S(x) = p_n(x)q(x) + r(x).$$

Note that the degree of $q(x)$ and $r(x)$ are less than n because if they were n or greater, the degree of $p_n(x)q(x) + r(x)$ would be greater than $2n - 1$.

Since $r(x)$ is of degree $n - 1$ or less, we can write $r(x)$ as a Lagrange interpolating polynomial using Legendre roots as the nodes. This yields

$$r(x) = \sum_{i=1}^n r(x_i) L_k(x).$$

Integrating both sides, we obtain

$$\begin{aligned} \int_{-1}^1 r(x) dx &= \int_{-1}^1 \sum_{i=1}^n r(x_i) L_k(x) \\ &= \sum_{i=1}^n r(x_i) \int_{-1}^1 L_k(x) dx. \end{aligned}$$

Notice that the right-hand side of this equation is simply the definition of Gaussian Quadrature, where $r(x_i)$ are the y -values at the Legendre roots, and $\int L_k(x) dx$ represent the coefficients. Thus, the integral of $r(x)$ is represented exactly by Gaussian Quadrature.

Note that by Theorem 3.5, $q(x)$ can be written as a linear combination of Legendre polynomials of degree less than n . So, consider the integral

$$\begin{aligned} \int_{-1}^1 S(x) dx &= \int_{-1}^1 (q(x)p_n(x) + r(x)) dx \\ &= \int_{-1}^1 \left(\sum_{i=0}^{n-1} a_i p_i(x) \right) p_n(x) dx + \int_{-1}^1 r(x) dx. \end{aligned}$$

Using the property of linearity we can write

$$\int_{-1}^1 S(x) dx = \sum_{i=0}^{n-1} \int_{-1}^1 a_i p_i(x) p_n(x) dx + \int_{-1}^1 r(x) dx.$$

Since $i \neq n$, $p_i(x)$ and $p_n(x)$ are orthogonal, and their product is equal to 0. It follows

that

$$\int_{-1}^1 S(x)dx = \int_{-1}^1 r(x)dx.$$

Note that if we let $x = x_i$, where x_i are the Legendre roots of $p_n(x)$, then

$$S(x_i) = p_n(x_i)q(x_i) + r(x_i)$$

$$S(x_i) = (0)(q(x_i)) + r(x_i)$$

$$S(x_i) = r(x_i),$$

since plugging the Legendre roots into the Legendre polynomial p_n produces 0. So, recognize that $S(x)$ and $r(x)$ agree at the Legendre roots. Therefore, their Gaussian Quadrature approximations are the same since $\int_{-1}^1 S(x)dx = \int_{-1}^1 r(x)dx$ and since Gaussian Quadrature is exact for $r(x)$, it must also be exact for $S(x)$. Thus, Gaussian Quadrature has a degree of precision $2n - 1$. \square

This theorem is important because it tells us that Gaussian Quadrature has a degree of precision that is twice as accurate as other numerical integration methods like the Newton-Cotes methods. However, we proved this fact only on the interval $[-1, 1]$. Another helpful property of Gaussian Quadrature is that it can be implemented no matter the interval. In order to use Gaussian Quadrature to approximate an integral $\int_a^b f(x)dx$ over an interval $[a, b]$, we must translate the integral back to $[-1, 1]$ [3]. In other words, we must map $x = -1$ to $x = a$ and $x = 1$ to $x = b$ using

$$x = \frac{(b - a)t + b + a}{2}.$$

Note that

$$dx = \frac{b-a}{2} dt.$$

So,

$$\int_a^b f(x) dx = \int_{-1}^1 f\left(\frac{(b-a)t + b + a}{2}\right) \frac{b-a}{2} dt.$$

To show that a and b are being mapped to -1 and 1 respectively, we will plug -1 and 1 into

$$f(t) = \frac{(b-a)t + b + a}{2}.$$

First, when $t = -1$,

$$\begin{aligned} f(-1) &= \frac{(b-a)(-1) + b + a}{2} \\ &= \frac{a - b + b + a}{2} \\ &= \frac{2a}{2} = a. \end{aligned}$$

Next, when $t = 1$,

$$\begin{aligned} f(1) &= \frac{(b-a)(1) + b + a}{2} \\ &= \frac{b - a + b + a}{2} \\ &= \frac{2b}{2} = b. \end{aligned}$$

Thus, we have mapped a to -1 and b to 1 .

We can demonstrate this property of Gaussian Quadrature using an example.

Example 3.10. *Approximate the integral*

$$\int_1^2 \ln(x) dx,$$

using Gaussian Quadrature.

Implementing this translation rule, we have

$$\begin{aligned}\int_1^2 \ln(x) dx &= \int_{-1}^1 \left(\frac{(2-1)t + 2+1}{2} \right) \frac{2-1}{2} dt \\ &= \int_{-1}^1 \left(\frac{t+3}{2} \right) \left(\frac{1}{2} \right) dt.\end{aligned}$$

So, let

$$f(t) = \ln \left(\frac{t+3}{2} \right) \left(\frac{1}{2} \right).$$

We will use the Legendre roots and Gaussian Quadrature coefficients from Table 2 to obtain a solution. First, using $n = 2$, we have

$$\begin{aligned}c_1 f(x_1) + c_2 f(x_2) &= 1f(-\sqrt{1/3}) + 1f(\sqrt{1/3}) \\ &\approx 0.3865949441.\end{aligned}$$

Next, using $n = 3$, we have

$$\begin{aligned}c_1 f(x_1) + c_2 f(x_2) + c_3 f(x_3) &= \frac{5}{9}f(-\sqrt{3/5}) + \frac{8}{9}f(0) + \frac{5}{9}f(\sqrt{3/5}) \\ &\approx 0.3863004215.\end{aligned}$$

Lastly, using $n = 4$, we have

$$\begin{aligned}c_1 f(x_1) + c_2 f(x_2) + c_3 f(x_3) + c_4 f(x_4) \\ \approx 0.3862944969.\end{aligned}$$

Note that the correct value of $2\ln 2 - 1 \approx 0.3862943611$ [3]. Again, we observe that as the number of nodes increases, the accuracy of our approximation increases as well. Also

recognize that we were able to use Gaussian Quadrature over the interval $[1, 2]$ using the interval mapping rule.

4 Conclusion

In this paper, we discussed the importance of polynomial interpolation, particularly Lagrange interpolation, and how it is implemented in numerical integration techniques. Included in these quadrature methods are the Newton-Cotes formulas for the Trapezoid Rule and Simpson's Rule, which help us approximate definite integrals. However, through further exploration, we determined a much more accurate tool we can use to make these approximations: Gaussian Quadrature. Gaussian Quadrature is valuable because it aids in approximating integrals that are considered impossible to solve exactly in closed form. For example, we can approximate values of $\int_{-1}^1 e^{-x^2/2} dx$, which is an essential component of probability and statistics. An important element of Gaussian Quadrature, the Legendre polynomials, are the orthogonal set of polynomials that produce the roots we use to implement this technique. The reason Gaussian Quadrature approximations are so much more accurate compared to approximations from other integration methods is their degree of precision, $2n - 1$, which tells us that Gaussian Quadrature produces approximations that are twice as accurate compared to Newton-Cotes methods. Although Gaussian Quadrature is initially defined over the interval $[-1, 1]$, it can be implemented over any interval using the translation rule that we analyzed, demonstrating that Gaussian Quadrature is an efficient, accurate way to produce approximations.

References

- [1] J. Keesling. Gaussian Quadrature <https://people.clas.ufl.edu/kees/files/GaussianQuadrature.pdf> (updated January 25, 2012)
- [2] R. Burden, D. Faires, and A. Burden, *Numerical Analysis*, 10th ed. Cengage Learning, Boston, MA, 2016.
- [3] T. Sauer, *Numerical Analysis*, 3rd ed. Pearson, Hoboken, NJ, 2019.
- [4] V. Elvira, L. Martino, P. Closas. Importance Gaussian Quadrature <https://par.nsf.gov/servlets/purl/10230639> (updated May 24, 2021)
- [5] J. Stewart, *Single Variable Calculus*, 8th ed. Cengage Learning, Boston, MA, 2015.
- [6] Wikipedia contributors, Wikipedia, The Free Encyclopedia. Gaussian Quadrature <https://w.wiki/8Svf> (updated December 5, 2023).
- [7] R. Williams. Normal Distribution <https://www3.nd.edu/~rwilliam/stats1/x21.pdf> (Updated September 2, 2004).